

density $\rho(r)$ as

$$\rho(r_j) = \sum_{i=1}^N w_i \rho_i(r_j) \quad \text{where} \quad \rho_i(r) \equiv \int d^3\mathbf{v} f_i(E, L) \quad (4.202)$$

is the density produced by the family of orbits that have the given energy and total angular momentum, but all possible orientations of the orbital plane. For given values of $\rho(r_j)$ these equations define a linear programming problem for the weights in the same way that equation (4.200) does. In fact, the only difference between these equations is that in one case the galaxy is decomposed into individual orbits, and in the other a symmetry principle is used to group orbits into families within which all orbits must have the same weight, and then the galactic density is written as a sum of the densities contributed by each family.

Schwarzschild modeling has been extensively used to search for massive black holes at the centers of luminous spheroids (e.g. Richstone & Tremaine 1985; van der Marel et al. 1998; Gebhardt et al. 2003), the results of which are summarized by the correlation (1.27) between black-hole mass and spheroid velocity dispersion. In §4.9.1 we shall see why reliable black-hole masses can be obtained only with sophisticated dynamical modeling of the observational data. Schwarzschild modeling has also been used to model the large-scale dynamics of early-type galaxies, thus constraining the mass densities and orbital distributions in these systems (Cappellari et al. 2006 and §4.9.2). Unfortunately, the inference of confidence intervals on the values of model parameters, such as black-hole masses and mass-to-light ratios, when Schwarzschild's method is used to fit a model to observational data, proves to be a subtle matter (Magorrian 2006), and published values should be treated with some caution.

4.8 The Jeans and virial equations

In §4.1.2 we saw that comparisons between theoretical models and observational data often center on velocity moments of the DF, such as $\bar{\mathbf{v}}$ and $\overline{v_i v_j}$. Calculating moments is easy if one knows the DF, but finding a DF that is compatible with a given probability density distribution $\nu(\mathbf{x})$ is not straightforward, and even if a DF can be found, it is often not unique. Therefore in this section we discuss techniques for inferring moments from stellar densities without actually recovering the DF. Dejonghe (1986) gives an extensive discussion of this problem.

Integrating equation (4.11) over all velocities, we obtain

$$\int d^3\mathbf{v} \frac{\partial f}{\partial t} + \int d^3\mathbf{v} v_i \frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \int d^3\mathbf{v} \frac{\partial f}{\partial v_i} = 0, \quad (4.203)$$

where we have employed the summation convention (page 772). The range of velocities over which we are integrating does not depend on time, so the partial derivative $\partial/\partial t$ in the first term of this equation may be taken outside the integral. Similarly, since v_i does not depend on x_i , the partial derivative $\partial/\partial x_i$ in the second term of the equation may be taken outside the integral sign. Furthermore, the last term on the left side of the equation vanishes on application of the divergence theorem (eq. B.46), given that $f(\mathbf{x}, \mathbf{v}, t) = 0$ for sufficiently large $|\mathbf{v}|$, i.e., there are no stars that move infinitely fast. Recalling the definitions of the density ν (eq. 4.20) and the mean velocity $\bar{\mathbf{v}}$ (eq. 4.24b), we have that

$$\frac{\partial \nu}{\partial t} + \frac{\partial(\nu \bar{v}_i)}{\partial x_i} = 0. \quad (4.204)$$

Equation (4.204) differs from the continuity equation (F.3) only in that it describes conservation of probability rather than that of mass, and replaces the fluid velocity by the mean stellar velocity.

We now multiply equation (4.11) by v_j and integrate over all velocities, and obtain

$$\frac{\partial}{\partial t} \int d^3\mathbf{v} f v_j + \int d^3\mathbf{v} v_i v_j \frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \int d^3\mathbf{v} v_j \frac{\partial f}{\partial v_i} = 0. \quad (4.205)$$

The last term on the left side can be transformed by applying the divergence theorem, using the fact that f vanishes for large $|\mathbf{v}|$:

$$\int d^3\mathbf{v} v_j \frac{\partial f}{\partial v_i} = - \int d^3\mathbf{v} \frac{\partial v_j}{\partial v_i} f = - \int d^3\mathbf{v} \delta_{ij} f = -\delta_{ij} \nu. \quad (4.206)$$

Thus equation (4.205) may be rewritten

$$\frac{\partial(\nu \bar{v}_j)}{\partial t} + \frac{\partial(\nu \bar{v}_i \bar{v}_j)}{\partial x_i} + \nu \frac{\partial \Phi}{\partial x_j} = 0. \quad (4.207)$$

This can be put into a more familiar form by subtracting from it \bar{v}_j times the equation of continuity (4.204) to yield

$$\nu \frac{\partial \bar{v}_j}{\partial t} - \bar{v}_j \frac{\partial(\nu \bar{v}_i)}{\partial x_i} + \frac{\partial(\nu \bar{v}_i \bar{v}_j)}{\partial x_i} = -\nu \frac{\partial \Phi}{\partial x_j}, \quad (4.208)$$

and then using the definition (4.26) of the velocity-dispersion tensor to eliminate $\bar{v}_i \bar{v}_j$. The result is an analog of Euler's equation (F.7) of fluid flow;

$$\nu \frac{\partial \bar{v}_j}{\partial t} + \nu \bar{v}_i \frac{\partial \bar{v}_j}{\partial x_i} = -\nu \frac{\partial \Phi}{\partial x_j} - \frac{\partial(\nu \sigma_{ij}^2)}{\partial x_i}. \quad (4.209)$$

The left side and the first term on the right side of equation (4.209) differ from terms in the ordinary Euler equation only by the replacement of the mass density by the probability density, and by the replacement of the fluid velocity by the mean stellar velocity. The last term on the right side of equation (4.209) represents something akin to the pressure force $-\nabla p$. More exactly, $-\nu\sigma_{ij}^2$ is a **stress tensor** that describes an anisotropic pressure. Since equations (4.204) and (4.209) were first applied to stellar dynamics by Jeans (1919), we call them the **Jeans equations**.¹⁶

Equations (4.204) and (4.209) are valuable because they relate observationally accessible quantities, such as the streaming velocity, velocity dispersion, and so forth. However, this is an incomplete set of equations in the following sense. If we know the potential $\Phi(\mathbf{x}, t)$ and the density $\nu(\mathbf{x}, t)$, we have nine unknown functions—the three components of $\bar{\mathbf{v}}$ and the six independent components of the symmetric tensor σ^2 —but only four equations—the scalar continuity equation and the three components of Euler's equation. Thus we cannot solve for $\bar{\mathbf{v}}$ and σ^2 without additional information. The reader may argue that if we multiply the collisionless Boltzmann equation (4.11) through by $v_i v_k$ and integrate over all velocities, we obtain a new set of differential equations for σ^2 which might supply the missing information. Unfortunately, these equations involve quantities like $\overline{v_i v_j v_k}$ for which we would require still further equations. Thus these additional equations are of no use unless we can in some way truncate or *close* this regression to ever higher moments of the velocity distribution. We shall find that closure is possible only in special circumstances, for example when the system is spherical and we know that its DF is ergodic, $f(H)$ (Box 4.3), or when the system is axisymmetric and its DF is of the form $f(H, L_z)$. The equations can also be closed for any Stäckel potential (van de Ven et al. 2003).

4.8.1 Jeans equations for spherical systems

To obtain the Jeans equations in spherical coordinates, we start from the collisionless Boltzmann equation in the form (4.14), which involves the canonical momenta

$$p_r = \dot{r} = v_r \quad ; \quad p_\theta = r^2 \dot{\theta} = r v_\theta \quad ; \quad p_\phi = r^2 \sin^2 \theta \dot{\phi} = r \sin \theta v_\phi. \quad (4.210)$$

We have

$$\int dp_r dp_\theta dp_\phi f = r^2 \sin \theta \int dv_r dv_\theta dv_\phi f = r^2 \sin \theta \nu. \quad (4.211)$$

We assume that the system is spherical and time-independent, so we can drop $\partial\Phi/\partial\theta$, $\partial\Phi/\partial\phi$, $\partial f/\partial t$ and $\partial f/\partial\phi$ from (4.14); we retain $\partial f/\partial\theta$ because any

¹⁶ They were originally derived by Maxwell, but he already has a set of equations named after him.

dependence of f on v_ϕ is likely to introduce θ -dependence through the last of equations (4.210) when v_ϕ is expressed in terms of p_ϕ . After simplification, equation (4.14) becomes

$$p_r \frac{\partial f}{\partial r} + \frac{p_\theta}{r^2} \frac{\partial f}{\partial \theta} - \left(\frac{d\Phi}{dr} - \frac{p_\theta^2}{r^3} - \frac{p_\phi^2}{r^3 \sin^2 \theta} \right) \frac{\partial f}{\partial p_r} + \frac{p_\phi^2 \cos \theta}{r^2 \sin^3 \theta} \frac{\partial f}{\partial p_\theta} = 0. \quad (4.212)$$

We now multiply by $p_r dp_r dp_\theta dp_\phi$ and integrate over all momenta. With equation (4.211) and similar results, and using the divergence theorem to eliminate derivatives with respect to the momenta, we find

$$\frac{\partial}{\partial r} (r^2 \sin \theta \overline{\nu p_r^2}) + \frac{\partial}{\partial \theta} (\sin \theta \overline{\nu p_r p_\theta}) + r^2 \sin \theta \nu \left(\frac{d\Phi}{dr} - \frac{\overline{p_\theta^2}}{r^3} - \frac{\overline{p_\phi^2}}{r^3 \sin^2 \theta} \right) = 0. \quad (4.213)$$

In any static spherical system, $\overline{p_r p_\theta} = r \overline{v_r v_\theta}$ must vanish because the DF is of the form $f(H, \mathbf{L})$, and is therefore an even function of v_r . Finally, dividing through by $r^2 \sin \theta$ and using equations (4.210) we obtain

$$\frac{d(\nu \overline{v_r^2})}{dr} + \nu \left(\frac{d\Phi}{dr} + \frac{2\overline{v_r^2} - \overline{v_\theta^2} - \overline{v_\phi^2}}{r} \right) = 0. \quad (4.214)$$

In terms of the anisotropy parameter of equation (4.61), equation (4.214) reads

$$\frac{d(\nu \overline{v_r^2})}{dr} + 2\frac{\beta}{r} \nu \overline{v_r^2} = -\nu \frac{d\Phi}{dr}. \quad (4.215)$$

Additional Jeans equations can be obtained by multiplying (4.212) by p_θ or p_ϕ , but these are not useful.

If the line-of-sight velocity dispersion has been measured as a function of radius, equation (4.215) can be used to constrain the radial dependence of β . The most direct approach is to assume a functional form for $\beta(r)$ and treat (4.215) as a first-order linear differential equation for $\nu \overline{v_r^2}$. The integrating factor is $\exp(2 \int dr \beta/r)$, so the solution can be written in closed form. Different choices of $\beta(r)$ yield different predictions for the line-of-sight velocity dispersion as a function of radius (see Problem 4.28), so β can be constrained by optimizing the fit between predictions obtained from (4.215) and the observed velocity-dispersion profile.

The case of constant non-zero β is particularly simple. Then the solution of (4.215) that satisfies the boundary condition $\lim_{r \rightarrow \infty} \overline{v_r^2} = 0$ is

$$\overline{v_r^2}(r) = \frac{1}{r^{2\beta} \nu(r)} \int_r^\infty dr' r'^{2\beta} \nu(r') \frac{d\Phi}{dr'}. \quad (4.216)$$

Effect of a central black hole on the observed velocity dispersion

We can use this equation to assess the impact of a central massive black hole

Box 4.3: Closure of the Jeans equations when the DF is ergodic

We have shown that the Jeans equations are not closed, in the sense that $\overline{v_r^2}$ and β cannot both be determined from ν and Φ . However, if the DF is known to be ergodic, $f(H)$, then $\beta = 0$ and $\overline{v_r^2}$ is determined by equation (4.216). Moreover, *all* of the n th-order velocity moments can be determined from $\overline{v_r^n}$ (Problem 4.29). A differential equation for $\overline{v_r^n}$ is obtained when we multiply (4.212) by $p_r^{n-1} dp_r dp_\theta dp_\phi$ and integrate over all momenta. For example, with $n = 4$ we find

$$\frac{d(\nu \overline{v_r^4})}{dr} = -3\nu \left(\frac{\overline{v_r^2} d\Phi}{dr} + \frac{\frac{2}{3}\overline{v_r^4} - \overline{v_r^2 v_\theta^2} - \overline{v_r^2 v_\phi^2}}{r} \right) \stackrel{\beta=0}{=} -3\nu \overline{v_r^2} \frac{d\Phi}{dr}, \quad (1)$$

where the first equality is valid for any spherical system and the second is obtained by assuming that $f = f(H)$ and using the relation $\overline{v_r^2 v_\theta^2} = \frac{1}{3}\overline{v_r^4}$ from (4.308). Once $\overline{v_r^2}(r)$ is known, we can solve (1) for $\overline{v_r^4}(r)$ and from that derive the other fourth-order moments. Then we can solve a similar equation for $\overline{v_r^6}(r)$ and so on up to whatever moment we desire. When moments up to order $n \sim 10$ have been determined, accurate predictions of LOSVDs can be made (Magorrian & Binney 1994). These predictions will be identical to those one could have obtained from Eddington's formula (4.46b) for f but will not enable us to check that f is non-negative.

on the host galaxy's velocity-dispersion profile. We assume that the galaxy has a constant mass-to-light ratio and is a Hernquist model of scale-length a —from equations (2.64) and (2.67), the density and potential are

$$\nu(r) = \frac{1}{2\pi a^2} \frac{1}{r(1+r/a)^3} \quad ; \quad \Phi(r) = -\frac{GM_g}{r+a} - \frac{G\mu M_g}{r}, \quad (4.217)$$

where $\mu = M_\bullet/M_g$ is the ratio of the black-hole mass M_\bullet to the galaxy mass M_g . Hence

$$\overline{v_r^2}(ax) = \frac{GM_g}{a} \frac{(1+x)^3}{x^{2\beta-1}} \int_x^\infty dx' \left(\frac{x'^{2\beta-1}}{(1+x')^5} + \frac{\mu x'^{2\beta-3}}{(1+x')^3} \right). \quad (4.218)$$

For integer values of 4β the integrals are elementary. For example with $y \equiv 1+x$ we have

$$\frac{a\overline{v_r^2}(ax)}{GM_g} = \begin{cases} 5(1+2\mu)x^2y^3 \ln(x/y) + \mu y^3(\frac{1}{3} - \frac{3}{2}x + 6x^2)/x \\ \quad + x^2[\frac{1}{4} + \frac{2}{3}y + \frac{3+\mu}{2}y^2 + 4(1+\mu)y^3]/y + xy^3 & (\beta = -\frac{1}{2}), \\ (1+6\mu)xy^3 \ln(y/x) - \mu y^3(3x - \frac{1}{2})/x \\ \quad - x[\frac{1}{4} + \frac{1}{3}y + \frac{1+\mu}{2}y^2 + (1+3\mu)y^3]/y & (\beta = 0), \\ 3\mu y^3 \ln(x/y) + 1/4y + \mu y(\frac{1}{2} + 2y + y^2/x) & (\beta = \frac{1}{2}). \end{cases} \quad (4.219)$$

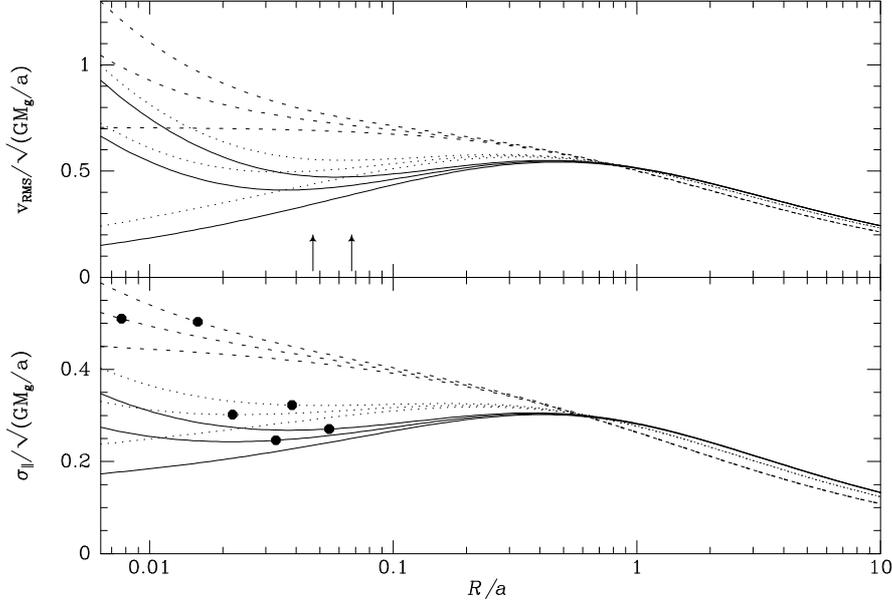


Figure 4.20 Velocity dispersion as a function of radius for three Hernquist models with a central black hole of mass 0, $0.002M_g$, or $0.004M_g$. The bottom panel shows line-of-sight dispersions, the top panel shows the RMS speed as a function of radius. The full curves are for tangential bias ($\beta = -0.5$), the dotted curves are for the isotropic model and the dashed curves are for radial bias ($\beta = 0.5$). The beads mark the radius of influence (eq. 4.220) of the black hole in each model, while the arrows mark the dynamical radius of the black hole, at which the interior mass of the galaxy equals the mass of the black hole.

The top panel of Figure 4.20 shows the RMS speed $v_{\text{RMS}} = (\overline{v_r^2} + \overline{v_\theta^2} + \overline{v_\phi^2})^{1/2}$ that follows from these formulae for $\mu = 0$, $\mu = 0.002$, and $\mu = 0.004$ (bottom to top). The full curves are for tangentially biased models, the dotted curves for isotropic models, and the dashed curves are for radially biased models. In each case the black hole causes the RMS speed to rise at small radii where its deep potential well speeds up the stars. The lower panel shows the associated line-of-sight dispersions. At small radii the upturn in σ_{\parallel} is much less sharp than that in v_{RMS} , because the signal from stars near the black hole is diluted by the light from foreground and background stars. Note also that the rise in dispersion associated with the black hole is difficult to distinguish from the rise in dispersion associated with radial anisotropy.

The black hole's **radius of influence** R_{infl} is defined to be the radius at which the Kepler speed due to the hole is equal to σ_{\parallel} . Quantitatively,

$$R_{\text{infl}} = \frac{GM_{\bullet}}{\sigma_{\parallel}^2(R_{\text{infl}})} = 11 \frac{M_{\bullet}}{10^8 \mathcal{M}_{\odot}} \left(\frac{\sigma_{\parallel}}{200 \text{ km s}^{-1}} \right)^{-2} \text{ pc}. \quad (4.220)$$

In Figure 4.20 R_{infl} is marked by a black dot on each relevant curve of $\sigma_{\parallel}(R)$.

It can be seen that at R_{infl} the black hole has increased σ_{\parallel} by a few percent, and that the contribution to σ_{\parallel} from the black hole increases fairly rapidly interior to R_{infl} .

Another measure of the radial extent of the black hole's influence is the **dynamical radius** r_g of the black hole, at which the gravitational forces from the black hole and the galaxy are equal, or, equivalently, the radius within which the galactic mass is equal to the black-hole mass. The dynamical radius r_g , unlike the radius of influence R_{infl} , depends only on the galaxy's mass distribution and not its kinematics. Orbits with apocenters inside r_g will be nearly Keplerian. The vertical arrows in Figure 4.20 mark r_g for the two black-hole masses considered. For the tangential models, r_g is only slightly larger than R_{infl} , while in the radially biased models, $r_g \sim 7R_{\text{infl}}$.

This discussion demonstrates that a major obstacle to detecting a central black hole using stellar kinematics is the degeneracy between the mass of the black hole and velocity anisotropy. This degeneracy can be lifted by obtaining data with higher spatial resolution than assumed in Figure 4.20. Alternatively, we can exploit the information contained in the entire LOSVD rather than just its second moment (§4.9.1).

4.8.2 Jeans equations for axisymmetric systems

For simplicity we assume that the system under study is in a steady state and axisymmetric so all derivatives with respect to t and ϕ vanish. With these assumptions (4.12) becomes

$$p_R \frac{\partial f}{\partial R} + p_z \frac{\partial f}{\partial z} - \left(\frac{\partial \Phi}{\partial R} - \frac{p_\phi^2}{R^3} \right) \frac{\partial f}{\partial p_R} - \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial p_z} = 0. \quad (4.221)$$

We multiply this equation by p_R , integrate over the momenta $p_R = v_R$, $p_\phi = Rv_\phi$, $p_z = v_z$, and then express the momenta in terms of velocities. In close analogy with our derivation of equation (4.215) we obtain

$$\frac{\partial(\overline{\nu v_R^2})}{\partial R} + \frac{\partial(\overline{\nu v_R v_z})}{\partial z} + \nu \left(\frac{\overline{v_R^2} - \overline{v_\phi^2}}{R} + \frac{\partial \Phi}{\partial R} \right) = 0. \quad (4.222a)$$

When we multiply (4.221) by p_z or p_ϕ rather than p_R , we obtain

$$\frac{1}{R} \frac{\partial(R \overline{\nu v_R v_z})}{\partial R} + \frac{\partial(\overline{\nu v_z^2})}{\partial z} + \nu \frac{\partial \Phi}{\partial z} = 0, \quad (4.222b)$$

$$\frac{1}{R^2} \frac{\partial(R^2 \overline{\nu v_R v_\phi})}{\partial R} + \frac{\partial(\overline{\nu v_z v_\phi})}{\partial z} = 0. \quad (4.222c)$$

If we assume that the density $\nu(R, z)$ and the confining potential $\Phi(R, z)$ are known, equations (4.222) constitute three constraints on the six second-order

Box 4.4: Two useful formulae

If we obtain ν by integrating $f(H, L_z)$ over all velocities, the resulting expression will depend on z only through $\Phi(R, z)$. In these circumstances it is advantageous to consider ν to be a function of (R, Φ) and equation (4.223) yields

$$\overline{\nu v_R^2}(R, z) = \int_{\Phi(R, z)}^0 d\Phi' \nu(R, \Phi') \quad (1)$$

Multiplying equation (4.224) by ν and using (1) we obtain

$$\overline{\nu v_\phi^2} = \frac{\partial}{\partial R} \left(R \int_{\Phi}^0 d\Phi' \nu(R, \Phi') \right) + \nu R \frac{\partial \Phi}{\partial R}.$$

In the first term on the right we carry the factor R inside the integral and then use the standard formula

$$\frac{d}{dx} \int_{f(x)}^0 dy g(x, y) = \int_{f(x)}^0 dy \frac{\partial g}{\partial x} - g(x, f) \frac{df}{dx}$$

to establish that

$$\overline{\nu v_\phi^2} = \int_{\Phi}^0 d\Phi' \frac{\partial}{\partial R} [R\nu(R, \Phi')]. \quad (2)$$

velocity moments. Thus, just as in the spherical case, the Jeans equations are not closed. However, if the DF is known to be of the form $f(H, L_z)$, the mixed moments in these equations will vanish, $\overline{v_R^2} = \overline{v_z^2}$, and the third equation becomes trivial. So we have two equations for two unknowns, and the system is closed. Specifically, (4.222b) can be integrated to yield (Nagai & Miyamoto 1976)

$$\overline{v_R^2}(R, z) = \overline{v_z^2}(R, z) = \frac{1}{\nu(R, z)} \int_z^\infty dz' \nu(R, z') \frac{\partial \Phi}{\partial z'}. \quad (4.223)$$

Now that $\overline{v_R^2}$ is known, we can obtain $\overline{v_\phi^2}$ from (4.224a):

$$\overline{v_\phi^2}(R, z) = \overline{v_R^2} + \frac{R}{\nu} \frac{\partial(\nu \overline{v_R^2})}{\partial R} + R \frac{\partial \Phi}{\partial R}. \quad (4.224)$$

Proceeding similarly with higher-order Jeans equations obtained by multiplying (4.221) by $p_z^{k+1} p_\phi^{n-k-2}$ for $k = 0, 1, \dots, n-2$ and by $p_R p_\phi^{n-2}$, we can relate all the n -order moments to either $\nu(R, z)$ or $\nu \overline{v_\phi^2}(R, z)$, depending on whether n is even or odd (Magorrian & Binney 1994). These moments will be identical to those one would have obtained by using the Hunter–Qian algorithm to calculate $f(\mathcal{E}, L_z)$ from the same data (§4.4.1).

(a) Asymmetric drift Figure 4.17 shows that in the solar neighborhood

the distribution of high-velocity stars is strongly asymmetric, in the sense that there are more stars lagging the LSR than leading it. We saw on page 325 that this phenomenon is nicely explained by the surface-density and velocity-dispersion gradients in the disk, and more quantitatively by Figure 4.16, but from equation (4.222a) we can easily recover its most important aspect, which is the asymmetric drift (page 326)

$$v_a \equiv v_c - \bar{v}_\phi, \quad (4.225)$$

where v_c is the circular speed in the solar neighborhood. We consider the values of v_a of a sequence of stellar populations, each with its own value of \bar{v}_R^2 .

We assume that the disk is in a steady state and is symmetric about its equator. Then, since the Sun lies close to the galactic equator, we may evaluate equation (4.222a) at $z = 0$. Since $\partial\nu/\partial z = 0$ by symmetry, we find

$$\frac{R}{\nu} \frac{\partial(\nu \bar{v}_R^2)}{\partial R} + R \frac{\partial(\bar{v}_R \bar{v}_z)}{\partial z} + \bar{v}_R^2 - \bar{v}_\phi^2 + R \frac{\partial\Phi}{\partial R} = 0 \quad (z = 0). \quad (4.226)$$

Using equation (4.26) to replace \bar{v}_ϕ^2 by the azimuthal velocity dispersion σ_ϕ^2 and using $R(\partial\Phi/\partial R) = v_c^2$, we obtain

$$\begin{aligned} \sigma_\phi^2 - \bar{v}_R^2 - \frac{R}{\nu} \frac{\partial(\nu \bar{v}_R^2)}{\partial R} - R \frac{\partial(\bar{v}_R \bar{v}_z)}{\partial z} &= v_c^2 - \bar{v}_\phi^2 \\ &= (v_c - \bar{v}_\phi)(v_c + \bar{v}_\phi) = v_a(2v_c - v_a). \end{aligned} \quad (4.227)$$

If we neglect v_a compared to $2v_c$, we obtain Stromberg's **asymmetric drift equation**

$$v_a \simeq \frac{\bar{v}_R^2}{2v_c} \left[\frac{\sigma_\phi^2}{\bar{v}_R^2} - 1 - \frac{\partial \ln(\nu \bar{v}_R^2)}{\partial \ln R} - \frac{R}{\bar{v}_R^2} \frac{\partial(\bar{v}_R \bar{v}_z)}{\partial z} \right]. \quad (4.228)$$

The value of the square bracket does not depend on the scale of the velocity-dispersion tensor $\bar{v}_i \bar{v}_j$, but only on the ratios of its components. So if two populations have similar density distributions $\nu(R, z)$ and velocity ellipsoids of the same shape and orientation, the square bracket will take the same value for both populations. Hence in this case $v_a \propto \bar{v}_R^2$. Figure 4.21 shows that a relationship of this type holds for main-sequence stars near the Sun. The horizontal axis shows the dispersions in the velocities normal to the line of sight for stars in each population. The vertical axis shows the average amount by which the stars lag the azimuthal motion of the Sun. Each data point is for one bin in stellar color $B - V$. The redder bins contain older stars, which have larger dispersions S because stars are gradually accelerated by fluctuations in the gravitational potential (§8.4). The intersection of the

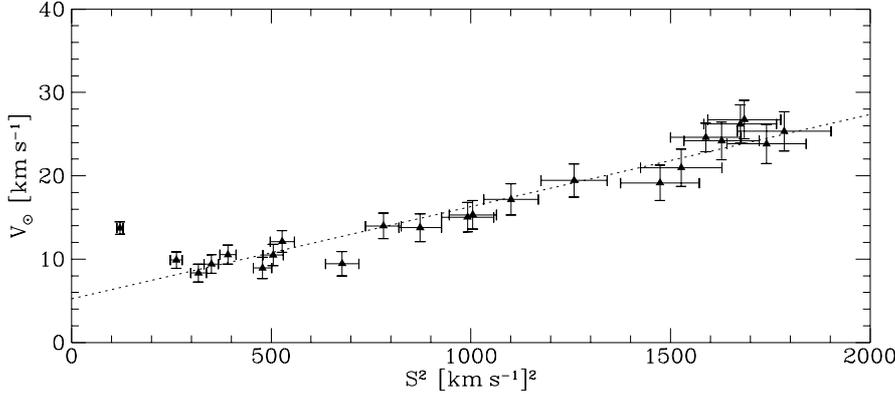


Figure 4.21 The asymmetric drift v_a for different stellar types is a linear function of the random velocity S^2 of each type. The vertical coordinate is actually $v_a + \tilde{v}_{\phi, \odot}$ where $\tilde{v}_{\phi, \odot}$ is the azimuthal velocity of the Sun relative to the LSR (after Dehnen & Binney 1998b).

best-fit line with $S = 0$, at $v_{\odot} = 5 \text{ km s}^{-1}$, represents the velocity of the Sun relative to the LSR.

It is interesting to compare the numerical value of the square bracket in equation (4.228) with the slope of the straight-line fit to the data in Figure 4.21. From BM Table 10.2 we adopt $\sigma_{\phi}^2/\overline{v_R^2} = 0.35$ and we assume that ν and $\overline{v_R^2}$ are both proportional to e^{-R/R_d} with $R_0/R_d = 3.2$ (Table 1.2)—this assumption regarding the radial dependence of the velocity dispersion is justified following equation (4.156). Then the bracket's first three terms sum to 5.8. The last term is problematic because its value depends on the orientation of the velocity ellipsoid at points just above the plane of our Galaxy, which is difficult to measure. Two extreme possibilities are that (i) the ellipsoid's principal axes are aligned with the coordinate directions of the (R, ϕ, z) system, and (ii) the principal axes are aligned with the coordinate directions of the (r, θ, ϕ) system centered on the galactic center. Orbit integrations (Binney & Spergel 1983) suggest that the truth lies nearly midway between these two possibilities. In the first case $\overline{v_R v_z}$ is independent of z and the term vanishes, and in the second $\overline{v_R v_z} \simeq (\overline{v_R^2} - \overline{v_z^2})(z/R)$ (see Problem 4.34) and the term contributes $-(1 - \overline{v_z^2}/\overline{v_R^2}) \simeq -0.8$. Averaging these values we estimate the value of the square bracket at 5.4 ± 0.4 , so $v_a \simeq \overline{v_R^2}/(82 \pm 6 \text{ km s}^{-1})$. From the data shown in Figure 4.21 one infers $v_a = \overline{v_R^2}/(80 \pm 5 \text{ km s}^{-1})$ in beautiful agreement with theory.

(b) Spheroidal components with isotropic velocity dispersion We know that if an axisymmetric system has a DF of the form $f(H, L_z)$ then two eigenvalues of the velocity-dispersion tensor σ^2 are equal (eq. 4.40). We now use the Jeans equations to predict the rotation rate of a spheroidal system in which all three eigenvalues of σ^2 are equal, that is, an isotropic rotator.

From the definition (4.26) of σ^2 and this assumption we have

$$\overline{v_\phi^2} = \overline{v_\phi^2} + \overline{\sigma_\phi^2} = \overline{v_\phi^2} + \overline{v_R^2}, \quad (4.229)$$

so equation (4.224) yields

$$\overline{v_\phi^2}(R, z) = R \frac{\partial \Phi}{\partial R} + \frac{R}{\nu} \frac{\partial(\nu \overline{v_R^2})}{\partial R}. \quad (4.230)$$

When we use equation (4.223) to eliminate $\overline{\nu v_R^2}$ we have

$$\overline{v_\phi^2}(R, z) = R \frac{\partial \Phi}{\partial R} + \frac{R}{\nu} \frac{\partial}{\partial R} \int_z^\infty dz' \nu(R, z') \frac{\partial \Phi}{\partial z'}. \quad (4.231)$$

Suppose both $\nu(R, z)$ and $\Phi(R, z)$ are constant on spheroids, which will be nearly true in many realistic cases. Then we can write $\nu(q_\nu^2 R^2 + \zeta)$ and $\Phi(q_\Phi^2 R^2 + \zeta)$ where $q_\nu < 1$ is the axis ratio of the isodensity surfaces, q_Φ is the axis ratio of the equipotentials, and $\zeta \equiv z^2$. Consequently, $\partial \nu / \partial R^2 = q_\nu^2 (\partial \nu / \partial \zeta)$ and $\partial \Phi / \partial R^2 = q_\Phi^2 (\partial \Phi / \partial \zeta)$. We convert the derivative in equation (4.231) into one with respect to R^2 , carry it under the integral sign, and use these relations to obtain

$$\begin{aligned} \overline{v_\phi^2}(R, z) &= R \frac{\partial \Phi}{\partial R} + \frac{2R^2}{\nu} \int_{z^2}^\infty d\zeta \left(q_\nu^2 \frac{\partial \nu}{\partial \zeta} \frac{\partial \Phi}{\partial \zeta} + q_\Phi^2 \nu \frac{\partial^2 \Phi}{\partial \zeta^2} \right) \\ &= R \frac{\partial \Phi}{\partial R} + (q_\nu^2 - q_\Phi^2) \frac{2R^2}{\nu} \int_{z^2}^\infty d\zeta \frac{\partial \nu}{\partial \zeta} \frac{\partial \Phi}{\partial \zeta} - 2R^2 q_\Phi^2 \frac{\partial \Phi}{\partial \zeta} \Big|_{z^2}, \end{aligned} \quad (4.232)$$

where the second equality is obtained by integrating by parts the term with the second derivative of Φ . We now observe that

$$2R^2 q_\Phi^2 \frac{\partial \Phi}{\partial \zeta} = 2R^2 \frac{\partial \Phi}{\partial R^2} = R \frac{\partial \Phi}{\partial R}. \quad (4.233)$$

Hence the last term on the right of (4.232) cancels the first term, and we have finally

$$\overline{v_\phi^2}(R, z) = (q_\nu^2 - q_\Phi^2) \frac{2R^2}{\nu} \int_{z^2}^\infty d\zeta \frac{\partial \nu}{\partial \zeta} \frac{\partial \Phi}{\partial \zeta}. \quad (4.234)$$

Since $\partial \Phi / \partial z > 0$ and $\partial \nu / \partial z < 0$, the integral is negative and $\overline{v_\phi} \propto \sqrt{q_\Phi^2 - q_\nu^2}$. If the isodensity surfaces coincide with the equipotentials, $q_\Phi = q_\nu$ and $\overline{v_\phi} = 0$, but normally the equipotentials are less flattened than the equidensity surfaces, and even a small excess in the flattening of the density distribution gives rise to appreciable rotation.

As an illustration of the use of equation (4.234), suppose Φ and ν are given by

$$\Phi = \frac{1}{2} v_0^2 \ln(R_c^2 + q_\Phi^2 R^2 + z^2) \quad ; \quad \nu = K(R_c^2 + q_\nu^2 R^2 + z^2)^{-3/2}, \quad (4.235)$$

where v_0 , R_c and K are constants. These functional forms are appropriate to the case of a galaxy that has a distribution of luminous matter consistent with a modified Hubble model (eq. 2.53) and an asymptotically flat circular-speed curve. Then the integral in equation (4.234) evaluates to

$$\begin{aligned} \int_{z^2}^{\infty} d\zeta \frac{\partial \nu}{\partial \zeta} \frac{\partial \Phi}{\partial \zeta} &= -\frac{3}{4} K v_0^2 \int_{z^2}^{\infty} \frac{d\zeta}{(R_c^2 + q_\Phi^2 R^2 + \zeta)(R_c^2 + q_\nu^2 R^2 + \zeta)^{5/2}} \\ &= -\frac{\frac{3}{2} K v_0^2}{(R_c^2 + q_\Phi^2 R^2 + z^2)^{5/2}} \frac{1}{\delta^4} \left(\frac{\sin^{-1} \delta}{\delta} + \frac{\frac{4}{3} \delta^2 - 1}{(1 - \delta^2)^{3/2}} \right). \end{aligned} \quad (4.236a)$$

where

$$\delta^2 \equiv \frac{(q_\Phi^2 - q_\nu^2) R^2}{R_c^2 + q_\Phi^2 R^2 + z^2}. \quad (4.236b)$$

To understand how this rather cumbersome formula works, we expand in powers of δ before substituting into equation (4.234) to obtain

$$\bar{v}_\phi^2(R, z) = \frac{3}{5} v_0^2 \left(\frac{R_c^2 + q_\nu^2 R^2 + z^2}{R_c^2 + q_\Phi^2 R^2 + z^2} \right)^{3/2} [\delta^2 + \frac{25}{14} \delta^4 + O(\delta^6)]. \quad (4.237)$$

At $R, z \ll R_c$, $\delta \propto R$ so $\bar{v}_\phi \propto R$ and there is solid-body rotation. Beyond R_c in the equatorial plane δ becomes independent of R and

$$\bar{v}_\phi/v_0 \rightarrow \sqrt{\frac{3}{5}(1 - q_\nu^2/q_\Phi^2)} (q_\nu/q_\Phi)^{3/2}. \quad (4.238)$$

Observations indicate that within the effective radius of a typical luminous galaxy, the mass distribution is dominated by stars (Gerhard et al. 2001; Cappellari et al. 2006). From §2.3.2 we know that the ellipticity $\epsilon_\nu \equiv 1 - q_\nu$ of the density distribution that generates the logarithmic potential (4.235) is $\simeq 3\epsilon_\Phi = 3(1 - q_\Phi)$. When we use this relation in equation (4.238), we find that $\bar{v}_\phi/v_0 = \sqrt{\frac{4}{5}\epsilon_\nu} + O(\epsilon_\nu^2)$. In Figure 4.14 the dotted curve shows the relationship $v/\sigma \propto \sqrt{\epsilon}$, and one sees that this proportionality provides a reasonable fit to the data for Evans models. It lies below the data for Rowley models, because these are not isotropic rotators. The filled circles, which show the data for low-luminosity spheroids, scatter around the dotted curve, although there is a tendency for the points to lie below the curve for $\epsilon \lesssim 0.5$ and above the curve at higher ellipticities. Thus these data suggest that low-luminosity spheroids are nearly isotropic rotators.

4.8.3 Virial equations

We obtained the Jeans equation (4.207) by multiplying the collisionless Boltzmann equation by v_j and integrating over all velocities. In this process an

equation in the six phase-space coordinates for a single scalar quantity f was reduced to partial differential equations for ν and the velocity moments in the three spatial coordinates. We now multiply equation (4.207) by x_k and integrate over all positions, thus converting these differential equations into a simple tensor equation relating global properties of the galaxy, such as total kinetic energy and mean-square streaming velocity.

We multiply equation (4.207) by Mx_k , where M is the total mass of the system. Then since the mass density is $\rho(\mathbf{x}) = M\nu(\mathbf{x})$, integrating over the spatial variables we find

$$\int d^3\mathbf{x} x_k \frac{\partial(\rho\bar{v}_j)}{\partial t} = - \int d^3\mathbf{x} x_k \frac{\partial(\rho\bar{v}_i\bar{v}_j)}{\partial x_i} - \int d^3\mathbf{x} \rho x_k \frac{\partial\Phi}{\partial x_j}. \quad (4.239)$$

The second term on the right side is the potential-energy tensor \mathbf{W} (eq. 2.19). The first term on the right side of equation (4.239) can be rewritten with the aid of the divergence theorem (B.45):

$$\int d^3\mathbf{x} x_k \frac{\partial(\rho\bar{v}_i\bar{v}_j)}{\partial x_i} = - \int d^3\mathbf{x} \delta_{ki} \rho\bar{v}_i\bar{v}_j = -2K_{kj}, \quad (4.240a)$$

where we have assumed that ρ vanishes at large radii and have defined the **kinetic-energy tensor**

$$K_{jk} \equiv \frac{1}{2} \int d^3\mathbf{x} \rho\bar{v}_j\bar{v}_k. \quad (4.240b)$$

With the help of equation (4.26) we split \mathbf{K} up into the contributions from ordered and random motion:

$$K_{jk} = T_{jk} + \frac{1}{2}\Pi_{jk}, \quad (4.241a)$$

where

$$T_{jk} \equiv \frac{1}{2} \int d^3\mathbf{x} \rho\bar{v}_j\bar{v}_k \quad ; \quad \Pi_{jk} \equiv \int d^3\mathbf{x} \rho\sigma_{jk}^2. \quad (4.241b)$$

The derivative with respect to time in equation (4.239) may be taken outside the integral sign because x_k does not depend on time. Finally, averaging the (k, j) and the (j, k) components of equation (4.239), we obtain

$$\frac{1}{2} \frac{d}{dt} \int d^3\mathbf{x} \rho (x_k\bar{v}_j + x_j\bar{v}_k) = 2T_{jk} + \Pi_{jk} + W_{jk}. \quad (4.242)$$

Here we have exploited the symmetry under exchange of indices of \mathbf{T} , $\mathbf{\Pi}$ (see eq. 4.241b) and \mathbf{W} (see eq. 2.22).

The left side of equation (4.242) may be brought to a more intuitive form if we define the tensor¹⁷ \mathbf{I} by

$$I_{jk} \equiv \int d^3\mathbf{x} \rho x_j x_k. \quad (4.243)$$

Differentiating \mathbf{I} with respect to time, we have

$$\frac{dI_{jk}}{dt} = \int d^3\mathbf{x} \frac{\partial \rho}{\partial t} x_j x_k. \quad (4.244)$$

With the continuity equation (4.204), the right side of this equation becomes

$$- \int d^3\mathbf{x} \frac{\partial(\rho \bar{v}_i)}{\partial x_i} x_j x_k = \int d^3\mathbf{x} \rho \bar{v}_i (x_k \delta_{ji} + x_j \delta_{ki}), \quad (4.245)$$

where the equality follows by an application of the divergence theorem. Substituting this expression back into equation (4.244) yields

$$\frac{dI_{jk}}{dt} = \int d^3\mathbf{x} \rho (x_k \bar{v}_j + x_j \bar{v}_k). \quad (4.246)$$

We now combine equations (4.242) and (4.246) to obtain the **tensor virial theorem**:

$$\frac{1}{2} \frac{d^2 I_{jk}}{dt^2} = 2T_{jk} + \Pi_{jk} + W_{jk}. \quad (4.247)$$

Equation (4.247) enables us to relate the gross kinematic and morphological properties of galaxies.¹⁸ In many applications the left side is simply zero since the galaxy is time-independent.

(a) Scalar virial theorem The trace of the potential-energy tensor is the system's total potential energy W (eq. 2.23). Equations (4.240b) show that $K \equiv \text{trace}(\mathbf{T}) + \frac{1}{2} \text{trace}(\mathbf{\Pi})$ is the total kinetic energy of the system. Thus, if the system is in a steady state, $\dot{\mathbf{I}} = 0$, and the trace of equation (4.247) becomes

$$2K + W = 0. \quad (4.248)$$

Equation (4.248) is a statement of the **scalar virial theorem**.¹⁹ The kinetic energy of a stellar system with mass M is just $K = \frac{1}{2} M \langle v^2 \rangle$, where $\langle v^2 \rangle$ is

¹⁷ The tensor defined by equation (4.243) is sometimes called the “moment of inertia tensor” but we reserve this name for the related tensor that is defined by equation (D.41).

¹⁸ Equation (4.247) has here been derived from the collisionless Boltzmann equation, which is only valid for a collisionless system, but we shall find in §7.2.1 that an analogous result is valid for any system of N mutually gravitating particles. However, it should be noted that (4.247) applies only to *self-gravitating* systems. Similar results may be derived for systems embedded in an externally generated gravitational field; see Problems 3.12 and 4.38.

¹⁹ First proved by R. Clausius in 1870; Clausius also defined the **virial** of a system of N particles as $\sum_i^N m_i \mathbf{r}_i \cdot \mathbf{v}_i$. The theorem was first applied to stellar systems by Eddington (1916a). Einstein (1921) used it to estimate the mass of globular clusters.

the mean-square speed of the system's stars. Hence the virial theorem states that

$$\langle v^2 \rangle = \frac{|W|}{M} = \frac{GM}{r_g}, \quad (4.249a)$$

where r_g is the gravitational radius defined by equation (2.42). One often wishes to estimate $\langle v^2 \rangle$ without going to the trouble of calculating r_g . Spitzer (1969) noted that in simple stellar systems the half-mass radius r_h , which is easily measured, is tightly correlated with r_g . For example, the Jaffe and Hernquist models (§2.2.2g) have $r_h/r_g = \frac{1}{2}$ and 0.402, respectively, while for spherical galaxies that have radius-independent mass-to-light ratios and satisfy the Sérsic law (1.17) in projection, r_h/r_g ranges from 0.414 for $m = 2$ to 0.526 for $m = 6$ (Ciotti 1991). Moreover, we saw in §4.3.3c that along the sequence of King models r_h/r_g is confined to the interval (0.4, 0.51) (Figure 4.10). Hence, a useful approximation is

$$\langle v^2 \rangle = \frac{|W|}{M} \simeq 0.45 \frac{GM}{r_h}. \quad (4.249b)$$

If E is the energy of the system, we have from equation (4.248) that

$$E = K + W = -K = \frac{1}{2}W. \quad (4.250)$$

Thus if a system forms by collecting material together from a state of rest at infinity (in which state, $K = W = E = 0$), and then settles by any process into an equilibrium condition, it invests half of the gravitational energy that is released by the collapse in kinetic form, and in some way disposes of the other half in order to achieve a binding energy $E_b = -E$ equal to its kinetic energy. For example, suppose that our Galaxy formed by aggregating from an initial radius that was much larger than its present size. Then, since most of the galactic material is now moving at about $v_c \simeq 200 \text{ km s}^{-1}$, whether on circular orbits in the disk or on eccentric and highly inclined halo orbits, we have that $E_b = K \approx \frac{1}{2}M_g v_c^2$ of energy must have been released when the Galaxy formed, where M_g is the mass of the Galaxy. This argument suggests that as they form, galaxies radiate a fraction $\frac{1}{2}(v_c/c)^2 \simeq 3 \times 10^{-7}$ of their rest-mass energy.

(b) Spherical systems We may use the scalar virial theorem (4.248) to evaluate the mass-to-light ratio Υ of a non-rotating spherical galaxy under the assumption that Υ is independent of radius. We choose a coordinate system in which the line of sight to the galaxy center coincides with the x axis. Then the kinetic energy associated with motion in the x -direction is

$$K_{xx} = \frac{1}{2} \int d^3\mathbf{x} \rho \overline{v_x^2}. \quad (4.251)$$